Figure 1

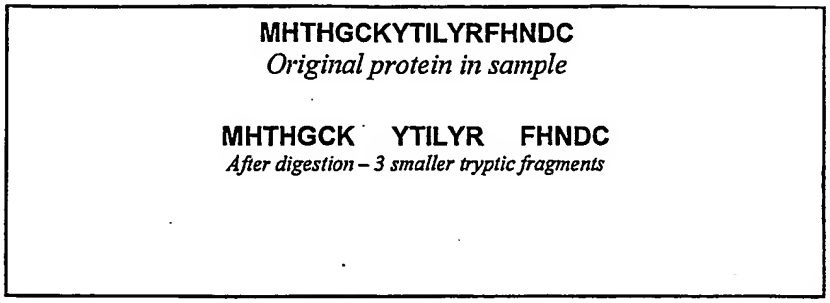

MHTHGCKYTILYRFHNDC
Original protein in sample
MHTHGCK YTILYR FHNDC
After digestion – 3 smaller tryptic fragments

Figure 2

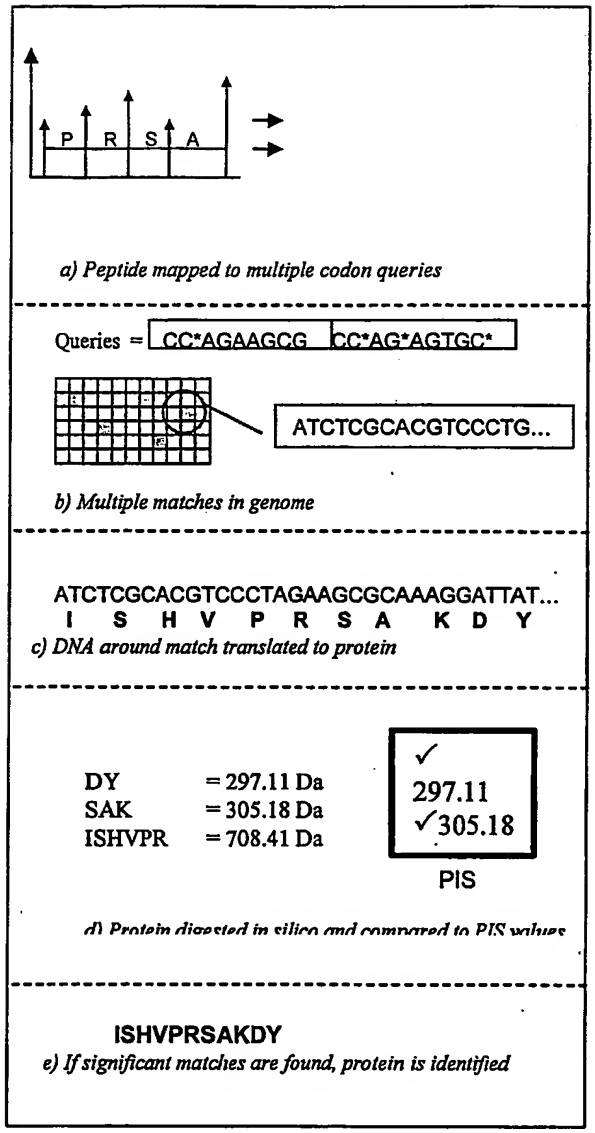

P R S A
a) Peptide mapped to multiple codon queries
Queries = CC*AGAAGCG CC*AG*AGTGC*
ATCTCGCACGTCCCTG...
b) Multiple matches in genome
ATCTCGCACGTCCCTAGAAGCGCAAAGGATTAT...
I S H V P R S A K D Y
c) DNA around match translated to protein
DY = 297.11 Da
SAK = 305.18 Da
ISHVPR = 708.41 Da
✓
297.11
✓305.18
PIS
d) Protein digested in silico and compared to PIS values
ISHVPRSAKDY
e) If significant matches are found, protein is identified

Figure 3

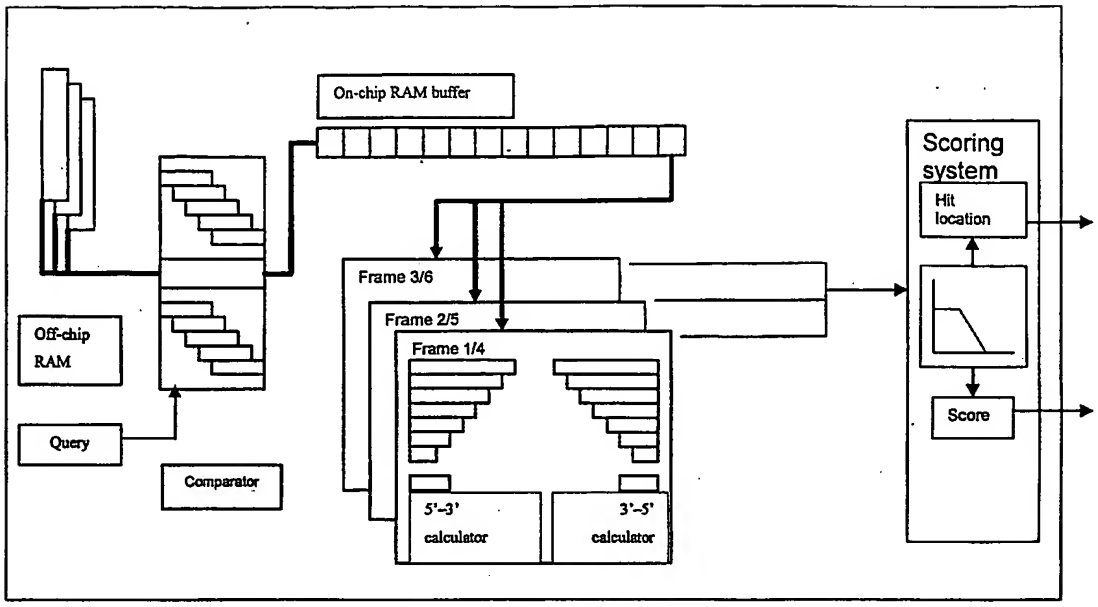

On-chip RAM buffer
Scoring system
Hit location
Frame 3/6
Frame 2/5
Frame 1/4
Off-chip RAM
Query
Comparator
5'-3' calculator
3'-5' calculator
Score

Figure 4

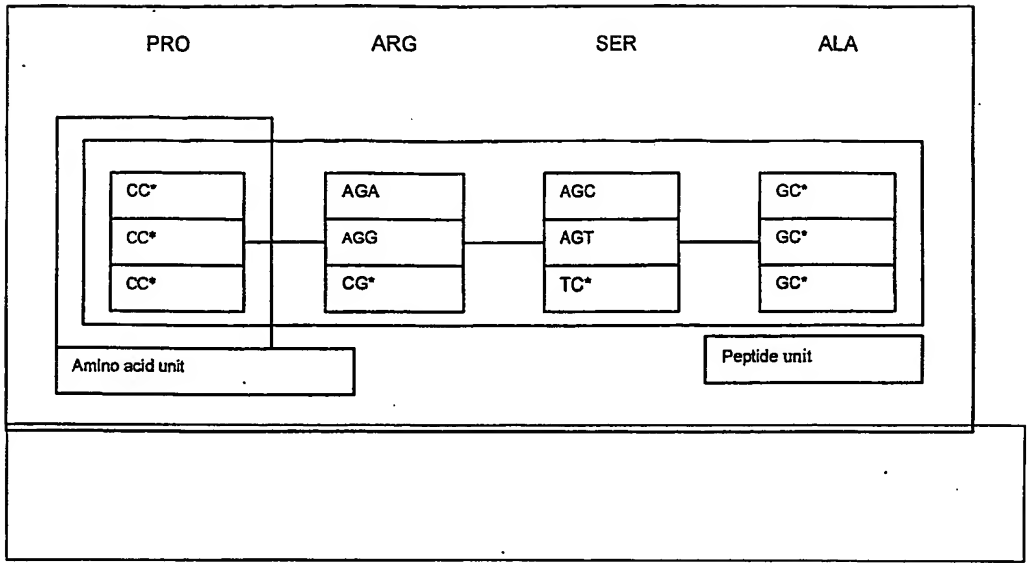

PRO
ARG
SER
ALA
CC*
CC*
CC*
AGA
AGG
CG*
AGC
AGT
TC*
GC*
GC*
GC*
Amino acid unit
Peptide unit

Figure 5

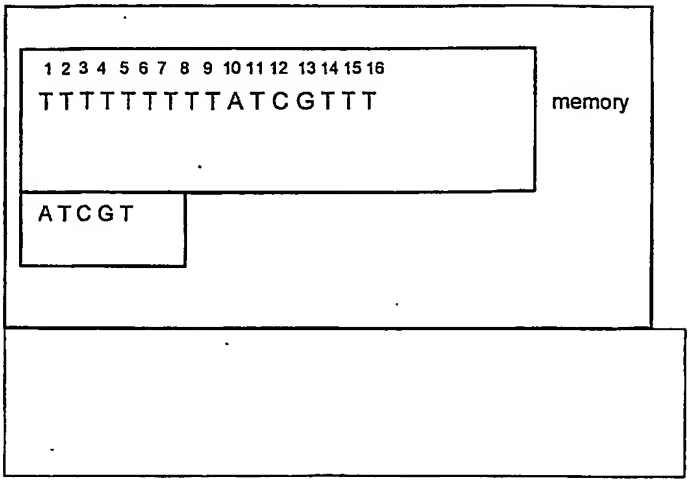

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
TTTTTTTTTATCGTTT
memory
ATCGT

Figure 6

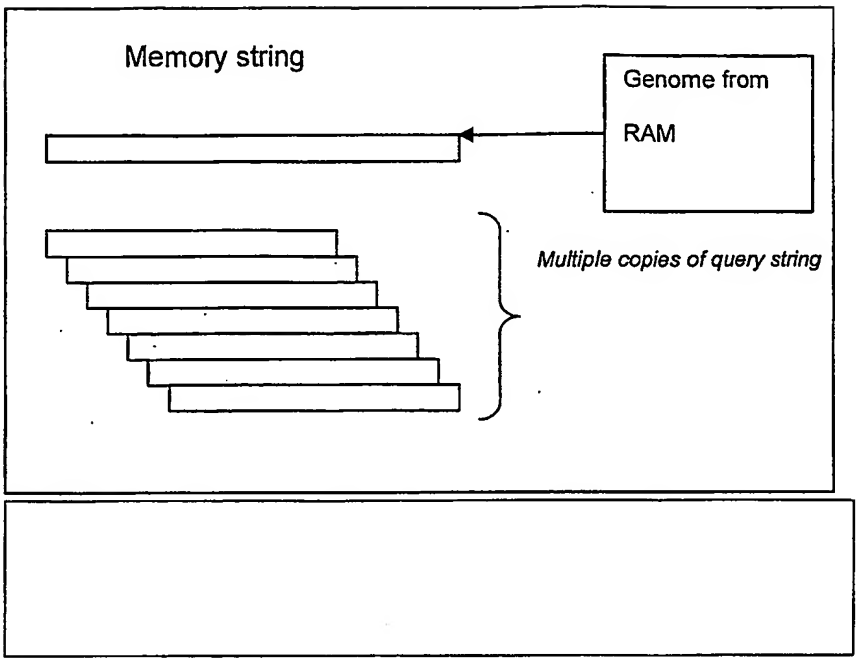
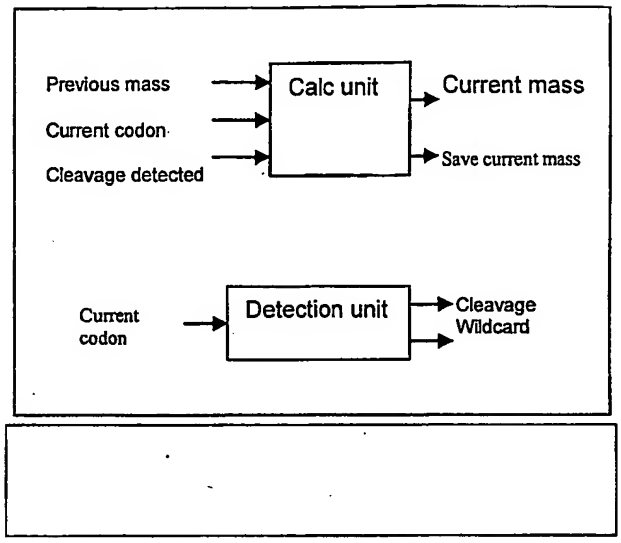
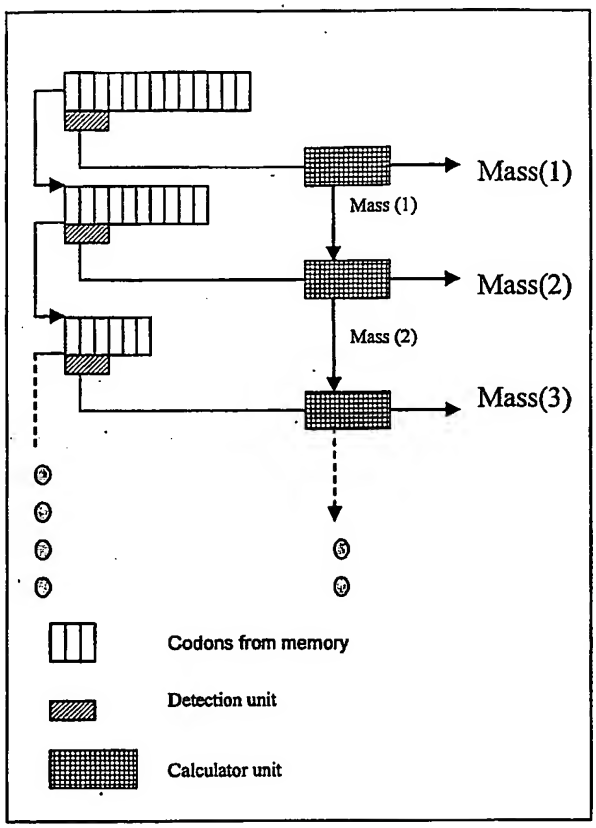

Memory string
Genome from
RAM
Multiple copies of query string

Figure 7


Previous mass
Current codon
Cleavage detected
Calc unit
Current mass
Save current mass
Current codon
Detection unit
Cleavage
Wildcard

8/38

Figure 8


Mass(1)
Mass (1)
Mass(2)
Mass (2)
Mass(3)
Codons from memory
Detection unit
Calculator unit

9/38

Figure 9

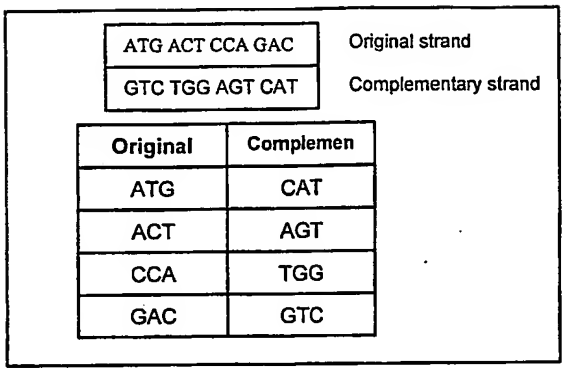

ATG ACT CCA GAC
Original strand
GTC TGG AGT CAT
Complementary strand
Original
Complemen
ATG
CAT
ACT
AGT
CCA
TGG
GAC
GTC

Figure 10

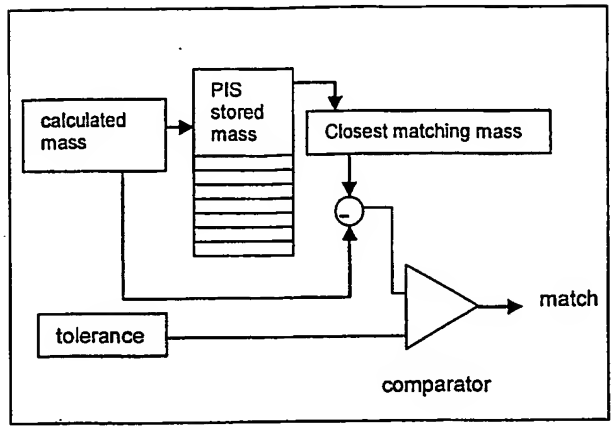
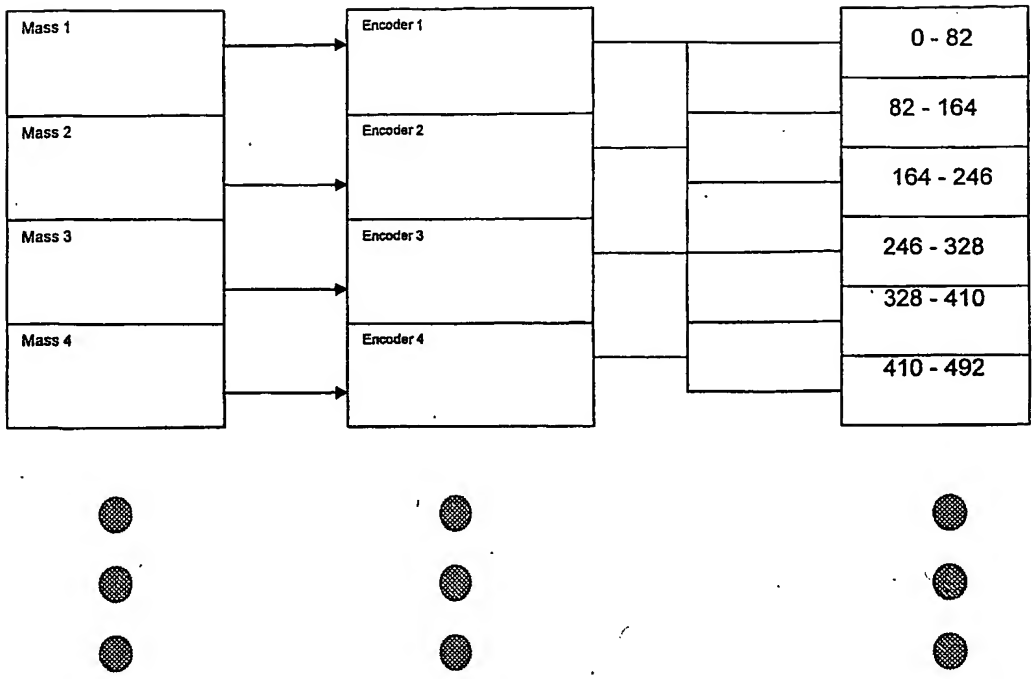
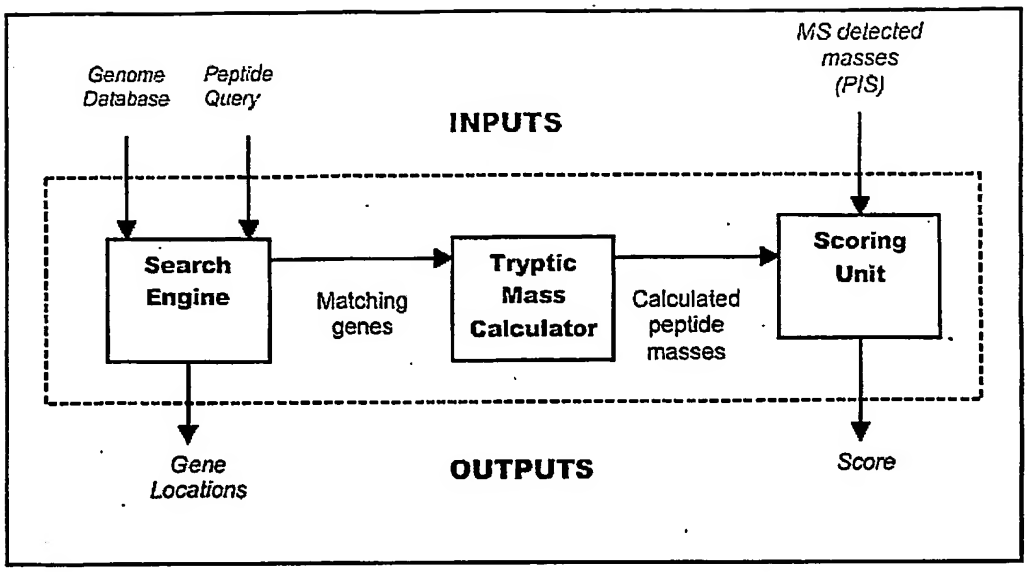

calculated mass
PIS stored mass
Closest matching mass
match
tolerance
comparator

Figure 11


Mass 1
Mass 2
Mass 3
Mass 4
Encoder 1
Encoder 2
Encoder 3
Encoder 4
0 - 82
82 - 164
164 - 246
246 - 328
328 - 410
410 - 492

12/38

Figure 12



13/38

Figure 13

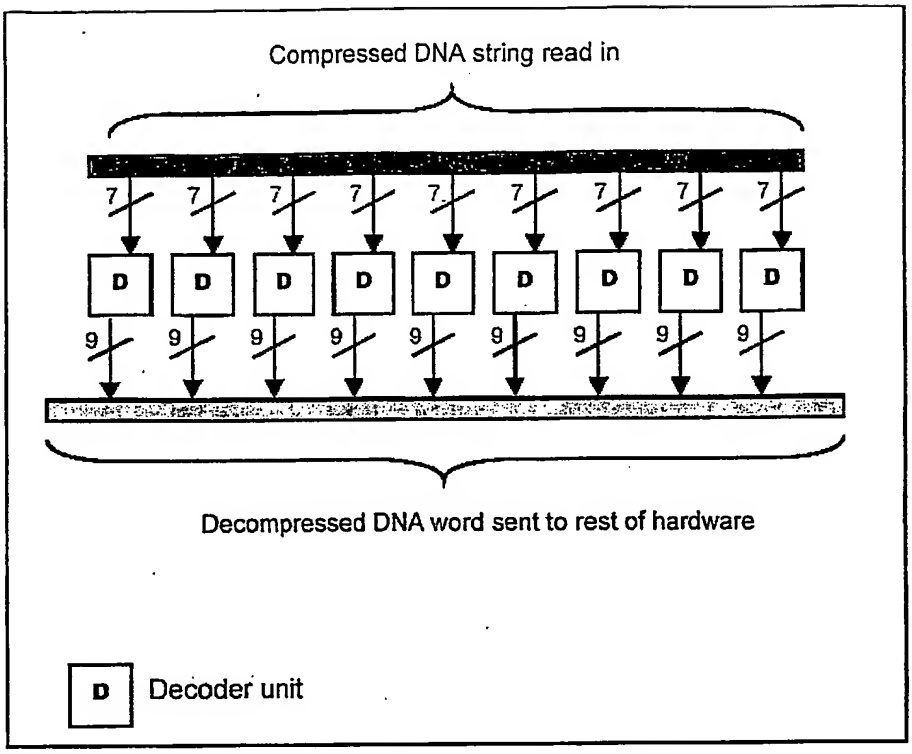
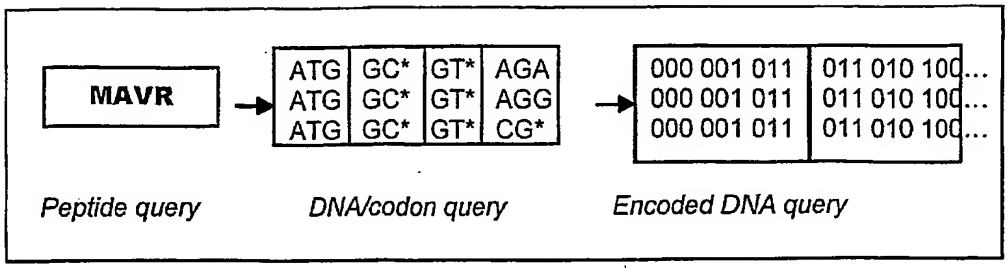


Figure 14


MAVR
ATG GC* GT* AGA
ATG GC* GT* AGG
ATG GC* GT* CG*
000 001 011 011 010 100...
000 001 011 011 010 100...
000 001 011 011 010 100...
Peptide query
DNA/codon query
Encoded DNA query

15/38

Figure 15

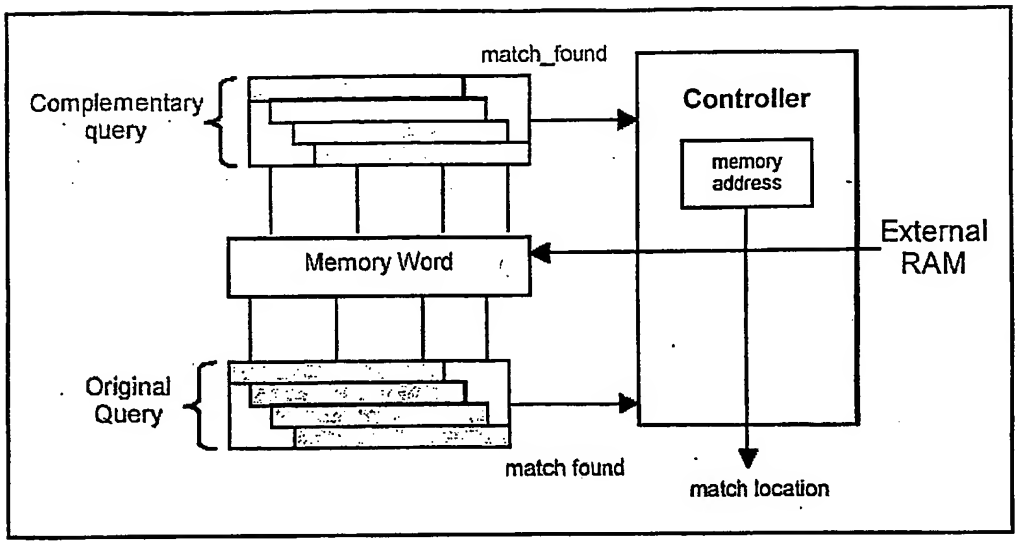


Figure 16

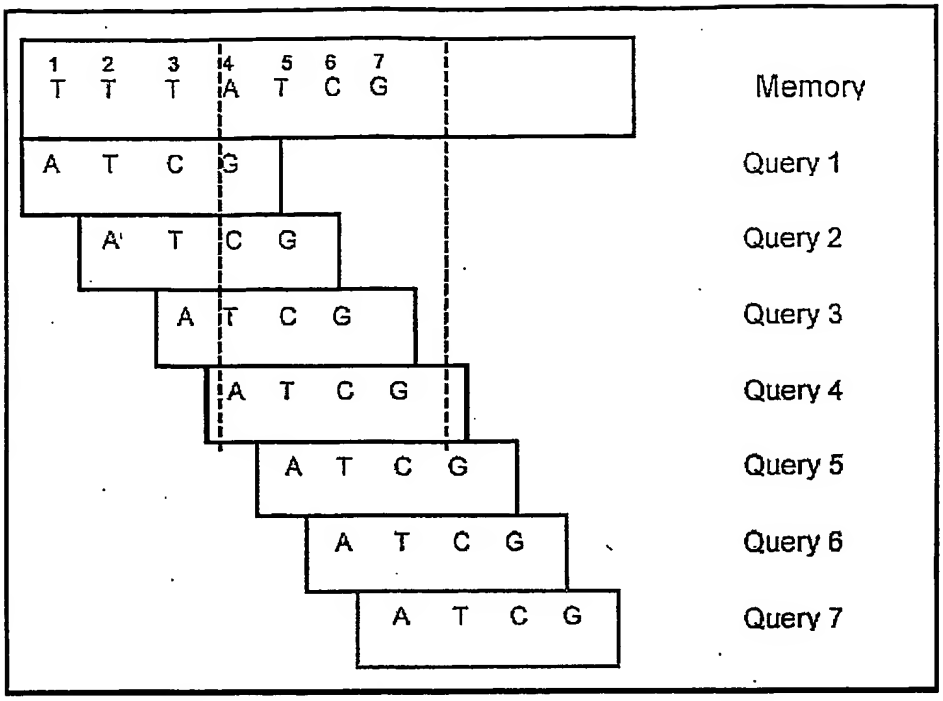

1 2 3 4 5 6 7
T T T A T C G
Memory
A T C G
Query 1
A T C G
Query 2
A T C G
Query 3
A T C G
Query 4
A T C G
Query 5
A T C G
Query 6
A T C G
Query 7

Figure 17

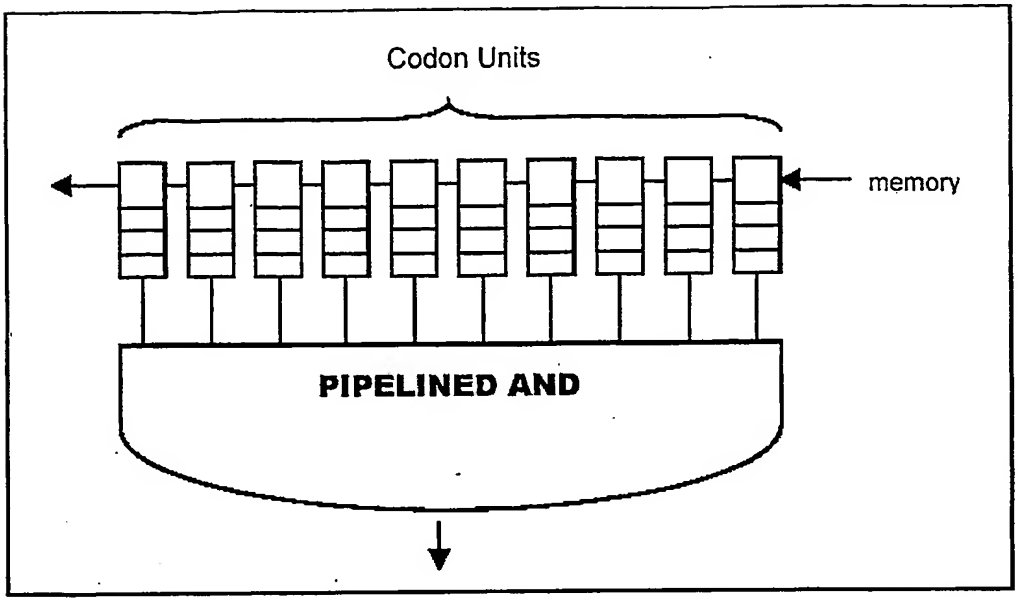


Figure 18

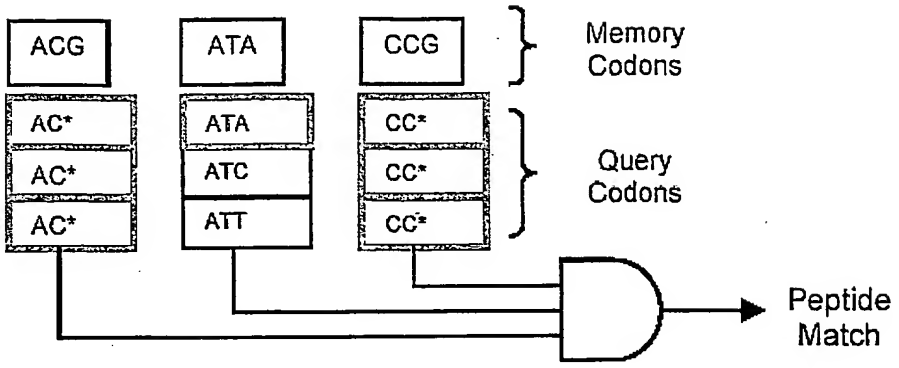
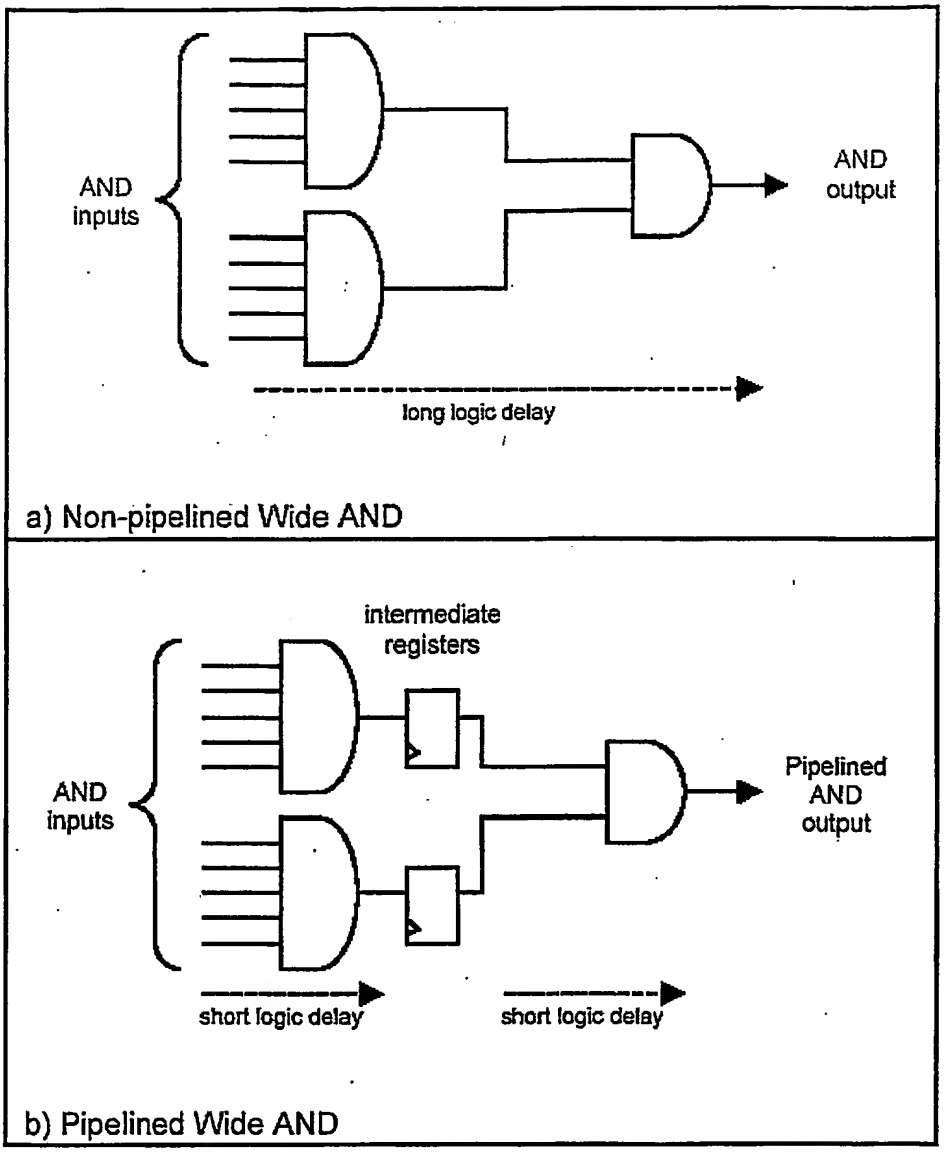
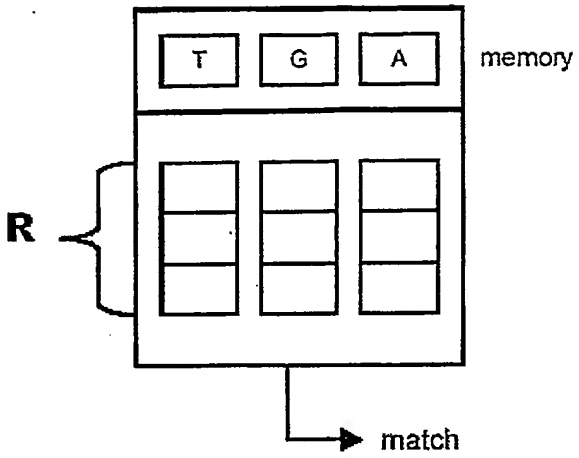

ACG
ATA
CCG
Memory Codons
AC*
AC*
AC*
ATA
ATC
ATT
CC*
CC*
CC*
Query Codons
Peptide Match

Figure 19


AND inputs
AND output
long logic delay
a) Non-pipelined Wide AND
intermediate registers
AND inputs
Pipelined AND output
short logic delay
short logic delay
b) Pipelined Wide AND

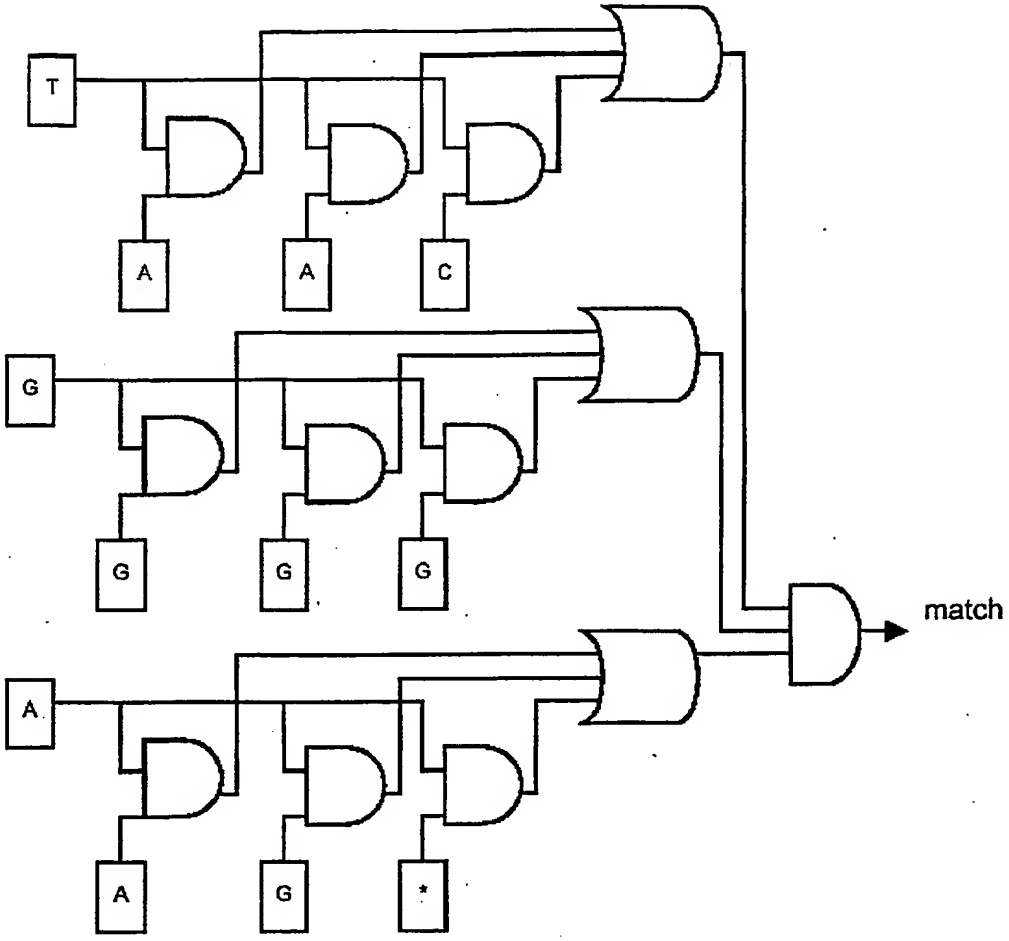
20/38

Figure 20



21/38

Figure 21



22/38

Figure 22

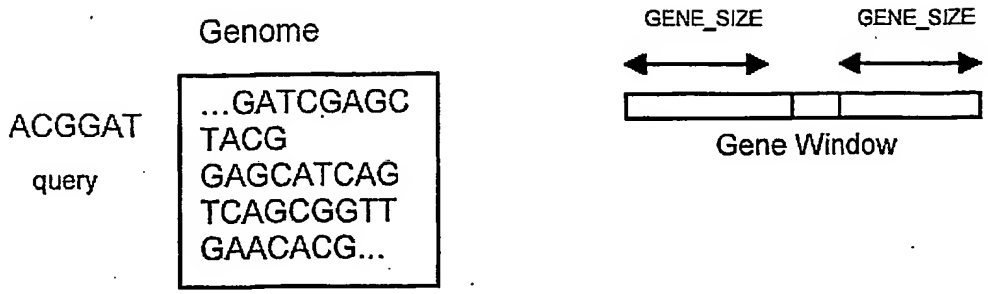


Figure 23

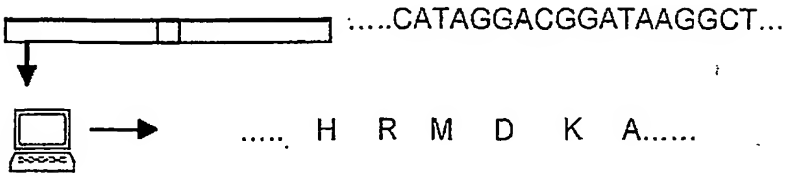
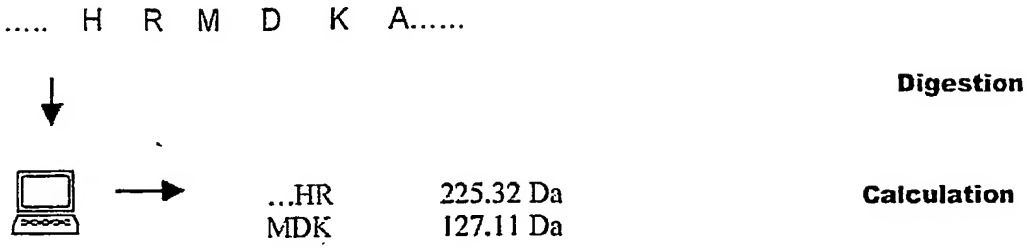
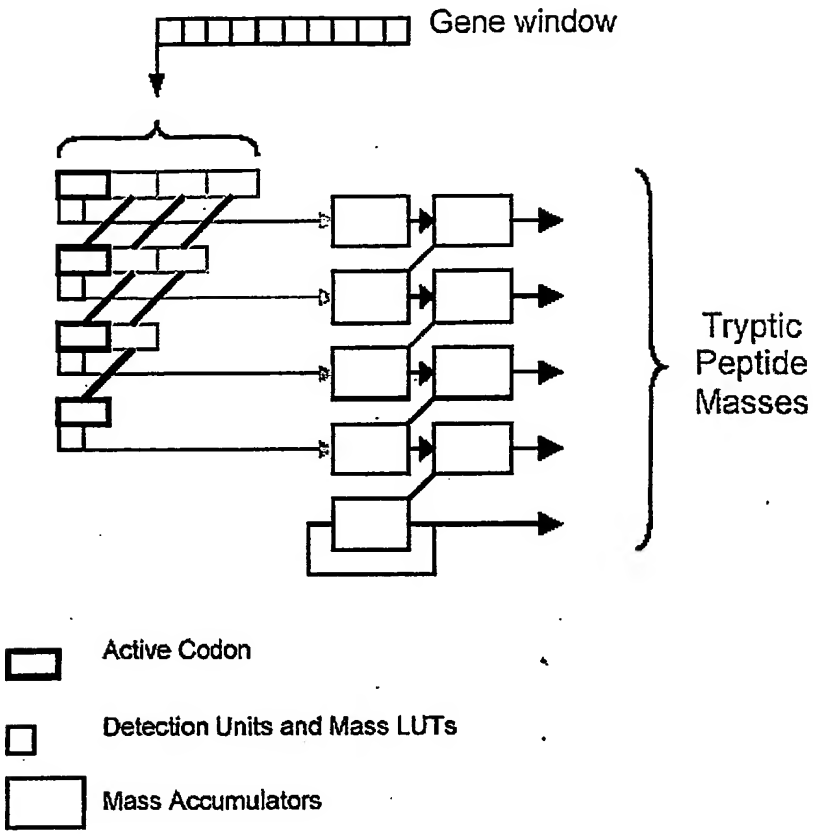


Figure 24



25/38

Figure 25



26/38

Figure 26

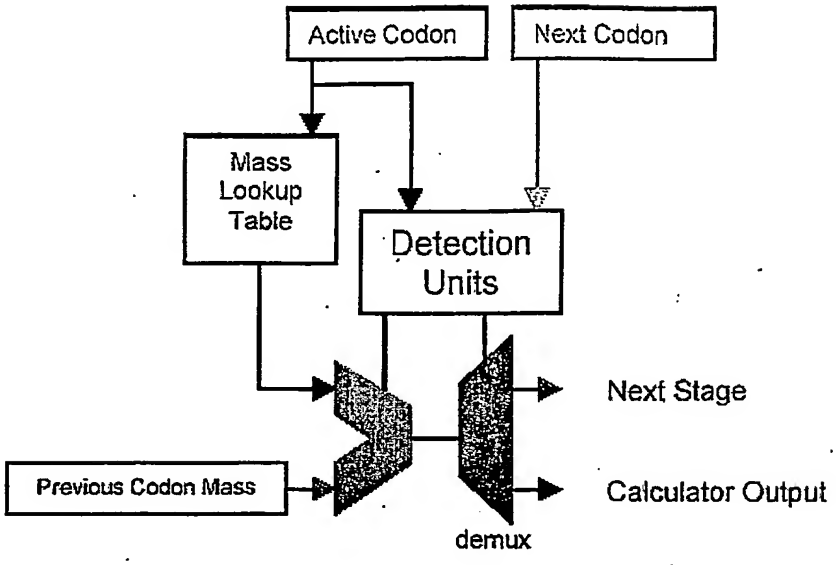
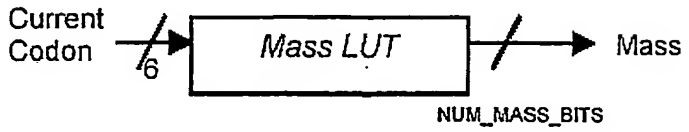
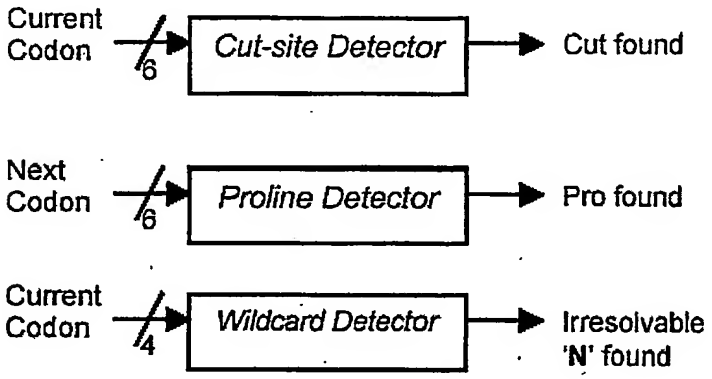


Figure 27



a) Mass Lookup Table(LUT)



b) Detection Units

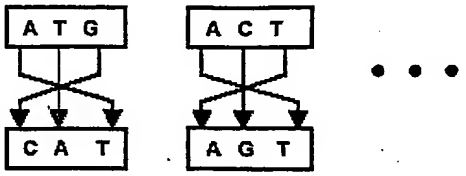
28/38

Figure 28

ATG ACT CCA GAC	Original strand
GTC TGG AGT CAT	Complementary strand

a) DNA Strand and its Complement

Original	Complement
ATG	CAT
ACT	AGT
CCA	TGG
GAC	GTC



a) Mapping codons to their complements

Figure 29

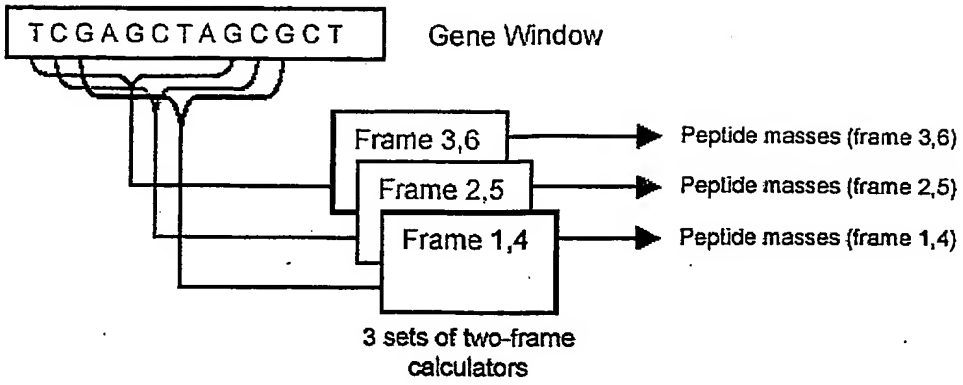
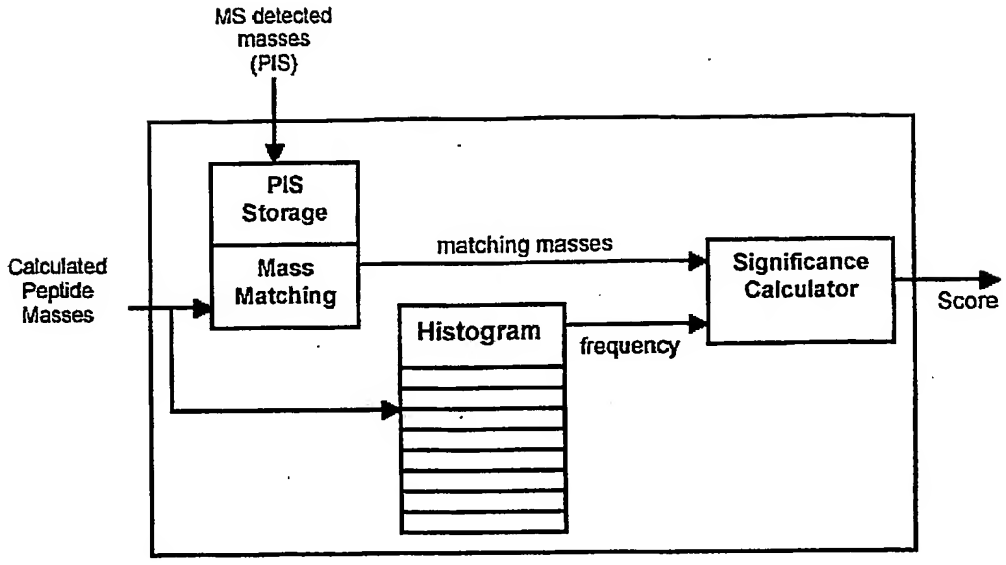
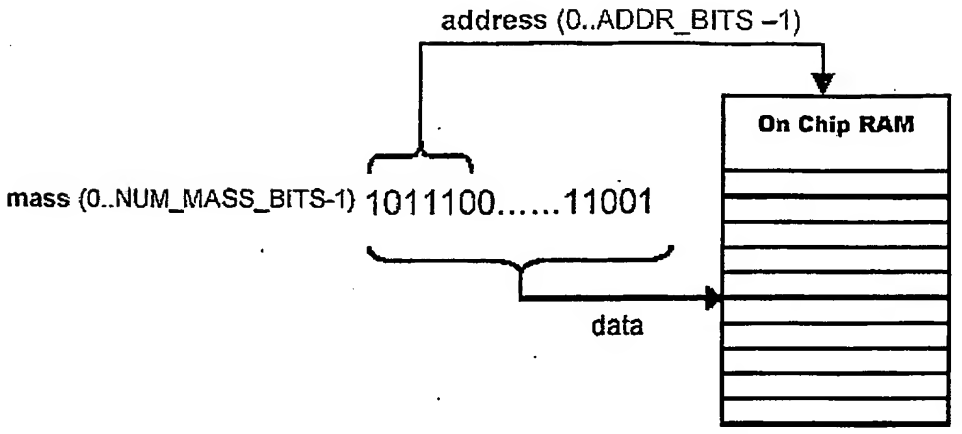

TCGAGCTAGCGCT
Gene Window
Frame 3,6
Frame 2,5
Frame 1,4
Peptide masses (frame 3,6)
Peptide masses (frame 2,5)
Peptide masses (frame 1,4)
3 sets of two-frame calculators

Figure 30


MS detected masses (PIS)
PIS Storage
Mass Matching
matching masses
Significance Calculator
Score
Calculated Peptide Masses
Histogram
frequency

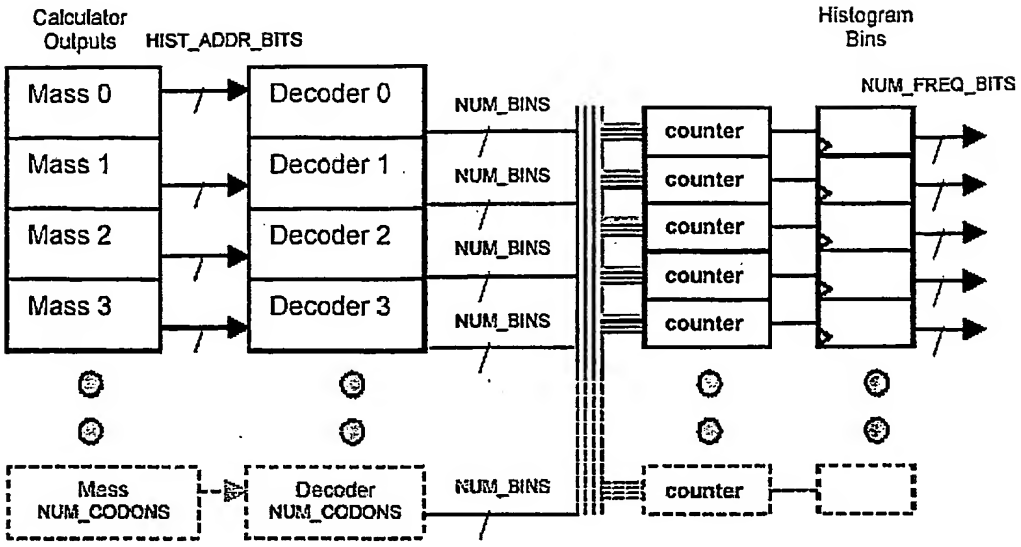
31/38

Figure 31



32/38

Figure 32



33/38

Figure 33

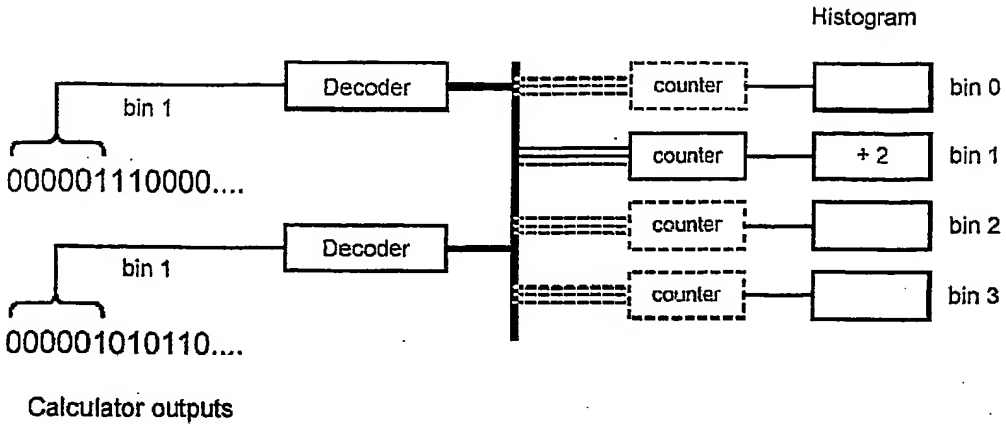


Figure 34

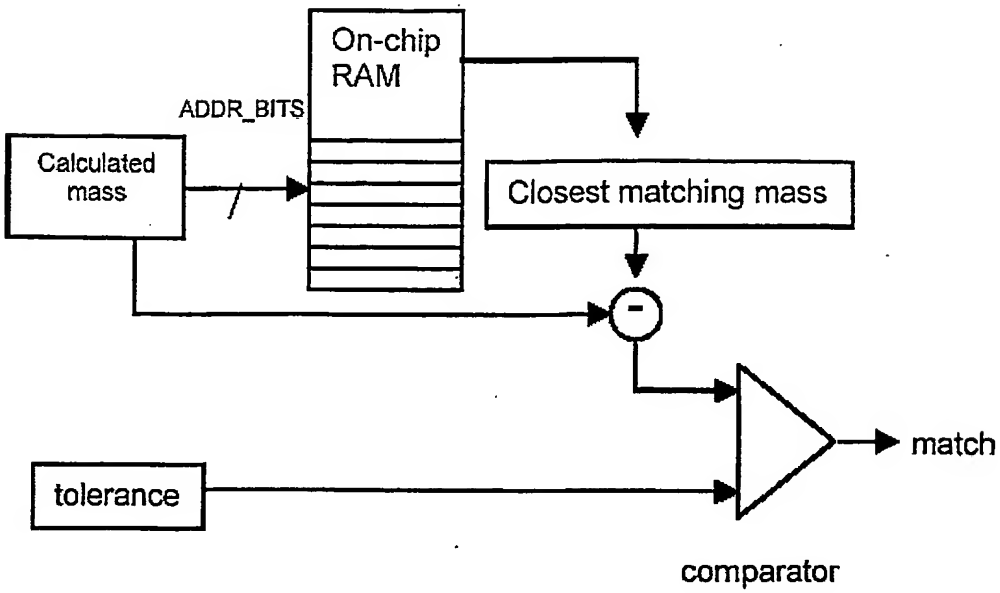
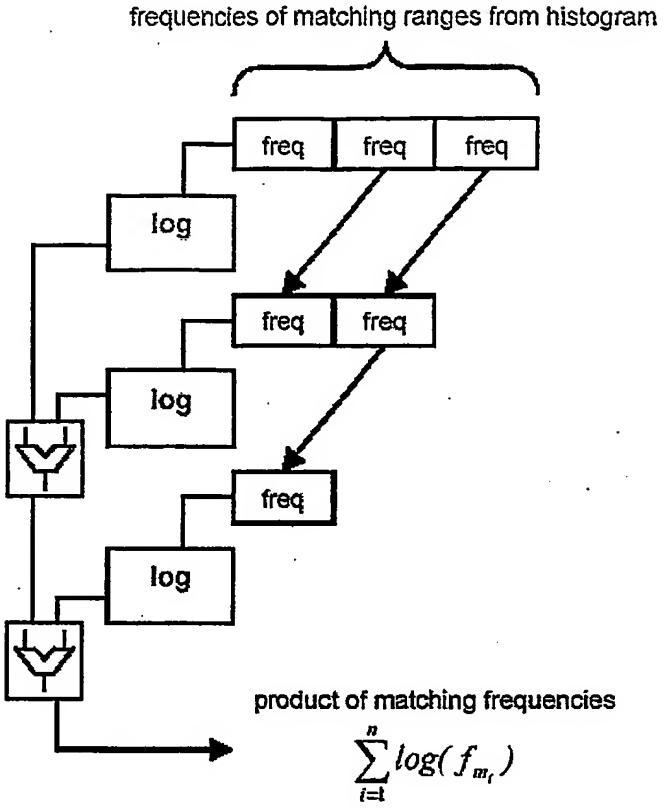
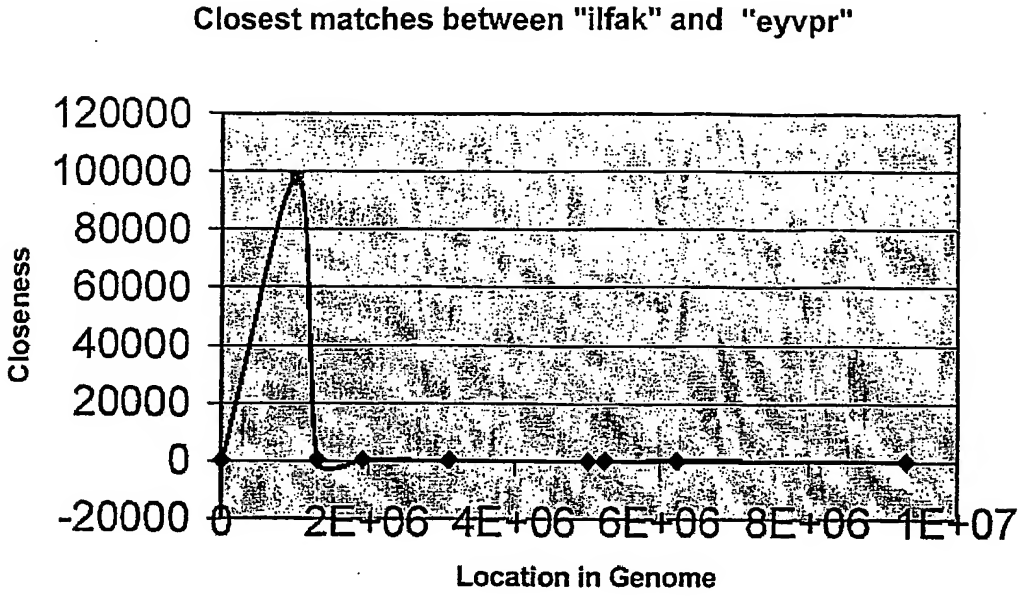

On-chip
RAM
ADDR_BITS
Calculated
mass
Closest matching mass
-
match
tolerance
comparator

Figure 35


frequencies of matching ranges from histogram
freq
freq
freq
log
freq
freq
log
freq
log
product of matching frequencies
$\sum_{i=1}^{n} log(f_{m_i})$

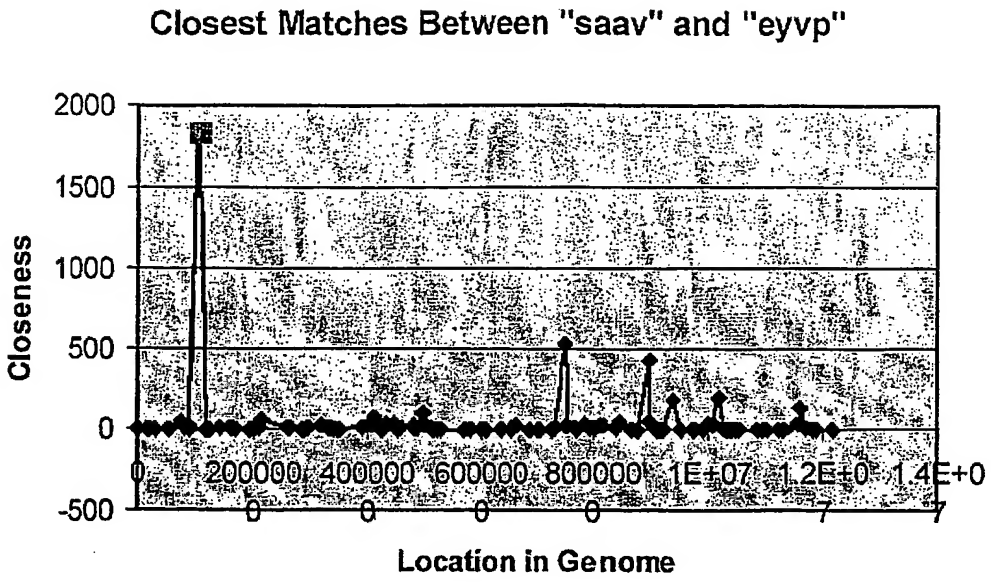
BEST AVAILABLE COPY

Figure 36



BEST AVAILABLE COPY

Figure 37



38/38

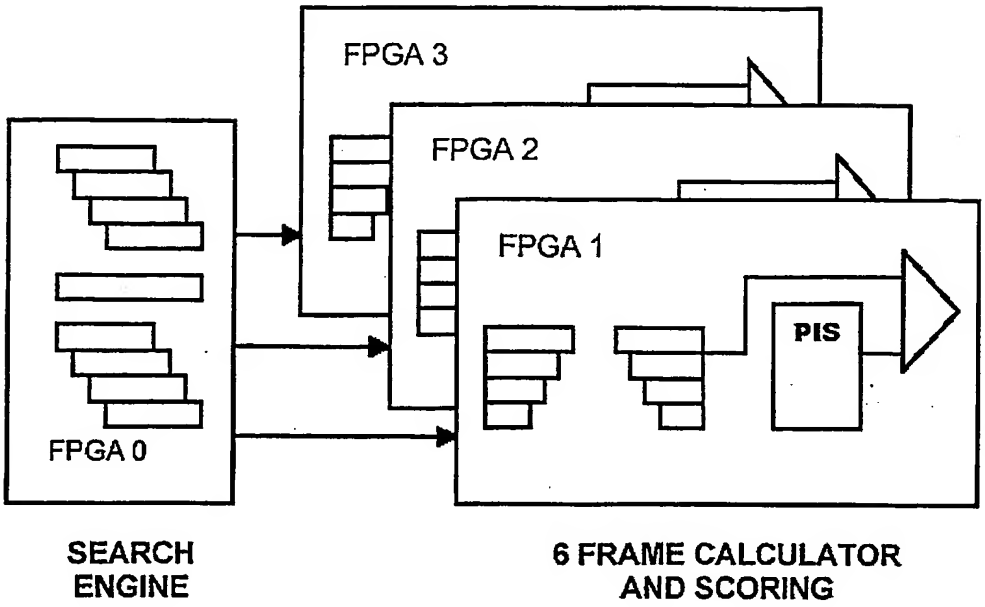



Figure 38